Video Games E-Commerce Sales Data Science Project

Datasets Description

Part I. VGChartz (Video Game Charts) Dataset

Our team's first dataset was selected from the publisher VGChartz (Video Game Charts): "a business intelligence and research firm … with an ever-expanding game database with over 55,000 titles" (Walton, *VGChartz*). Our team's dataset contains a list of video games with game sales figures greater than 100,000 copies.

Attributes to note are the 33 unique consoles, the years a game was released, the genre, the 580 unique publishers, the regional sales (NA, EU, JP, and other regions) in the millions, and the total worldwide sales in the millions.

When exploring the dataset, we observed that many of the existing games had their original and subsequent releases, which meant that franchise games were repeated because of multiple releases on different consoles. Our team wanted to evaluate the performance of these games not only by their sales figures but also by the gaming scores from critics and regular gamers.

```
Int64Index: 16325 entries, 0 to 16597
Data columns (total 10 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   Name          16325 non-null  object
 1   Platform      16325 non-null  object
 2   Year          16325 non-null  int32
 3   Genre         16325 non-null  object
 4   Publisher     16289 non-null  object
 5   NA_Sales      16325 non-null  float64
 6   EU_Sales      16325 non-null  float64
 7   JP_Sales      16325 non-null  float64
 8   Other_Sales   16325 non-null  float64
 9   Global_Sales  16325 non-null  float64
dtypes: float64(5), int32(1), object(4)
memory usage: 1.3+ MB
```

Part II. Metacritic Dataset

To incorporate the gaming scores into our exploration, our team's second dataset was selected from the Metacritic website: critics' consensus in one place with a single meta score and streamlining their user voting process (Editorial team, *Metacritic*). In our team's dataset, we were able to compile a list of critic and user scores of the best video games of all time.

Attributes to note are the 22 unique consoles, release dates spanning 1996 to 2017, meta critic score (from 0 to 100), and user critic score (from 0 to 10). When our team observed the meta-critic dataset, we noticed that there were much fewer observations.

```
RangeIndex: 18800 entries, 0 to 18799
Data columns (total 6 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   name          18800 non-null  object
 1   platform      18800 non-null  object
 2   release_date  18800 non-null  int64
 3   summary       18686 non-null  object
 4   meta_score    18800 non-null  int64
 5   user_review   18800 non-null  object
dtypes: int64(2), object(4)
memory usage: 881.4+ KB
```

This can be explained by the ‘All Platform’ filter in place when the data was retrieved from the Metacritic website so each game includes scores from their original release (unless the games were part of a franchise). Though our datasets may contain a multitude of missing values, our team was confident in their ability to analyze the correlation between video game sales and scores.

Merge Datasets

Our team pulled the VGChartz and Metacritic datasets from the Kaggle website and imported them into the notebook. We proceeded to measure the length of specific columns in each dataset for unique data values and checked for NaN values.

The columns of the second data frame were renamed to allow for an outer merge of two datasets - *vgsales.csv* and *all_games.csv* - calling the new data frame *df_combined*. Then, our team created a new data frame called *df_merged* grouping by name and aggregating mean row values of all numeric variables from *df_combined*.

	Name	Platform	Year	Genre	Publisher	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Summary	Critic_Score	User_Score
0	Wii Sports	Wii	2006	Sports	Nintendo	41.49	29.02	3.77	8.46	82.74	Tennis (1-4 players): Players grab the control...	76.0	8.1
1	Super Mario Bros.	NES	1985	Platform	Nintendo	29.08	3.58	6.81	0.77	40.24	NaN	NaN	NaN
2	Mario Kart Wii	Wii	2008	Racing	Nintendo	15.85	12.88	3.79	3.31	35.82	Mario Kart Wii comes with the intuitive Wii Wh...	82.0	8.4
3	Wii Sports Resort	Wii	2009	Sports	Nintendo	15.75	11.01	3.28	2.96	33.00	Wii Sports Resort is a collection of fun sport...	80.0	8.2
4	Pokemon Red/Pokemon Blue	GB	1996	Role-Playing	Nintendo	11.27	8.89	10.22	1.00	31.37	NaN	NaN	NaN
5	Tetris	GB	1989	Puzzle	Nintendo	23.20	2.26	4.22	0.58	30.26	NaN	NaN	NaN
6	New Super Mario Bros.	DS	2006	Platform	Nintendo	11.38	9.23	6.50	2.90	30.01	The first new 2D Mario platformer since Super ...	89.0	8.5
7	Wii Play	Wii	2006	Misc	Nintendo	14.03	9.20	2.93	2.85	29.02	NaN	NaN	NaN
8	New Super Mario Bros. Wii	Wii	2009	Platform	Nintendo	14.59	7.06	4.70	2.26	28.62	New Super Mario Bros. Wii offers a combination...	87.0	8.3
9	Duck Hunt	NES	1984	Shooter	Nintendo	26.93	0.63	0.28	0.47	28.31	NaN	NaN	NaN

We checked the length of the combined datasets to see if the data values were filtered by the group by function.

Data Cleaning

We cleaned *df_combined* by removing the rows with NaN values as some games were released in different years such as games being released on multiple consoles. Right after we cleared the NaN values, we filled NaN values for the Genre and Publisher categorical variables with the value ‘Unknown’ when the list of games was last updated.

Using the dropna command drops any NaN value from rows that are mostly missing sales and scores values as a result of the outer merge.

Once we cleaned the grouped data frame *df_merged* and this left rows of game copies that are missing sales and score values, we applied the dropna function again to drop all the NaN values and rename the columns to match the *df_combined* data frame.

	Name	Year	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	User_Score
0	.hack//Infection Part 1	2002.0	0.490000	0.380	0.260000	0.130000	1.270000	75.0	8.30
1	.hack//Mutation Part 2	2002.0	0.230000	0.180	0.200000	0.060000	0.680000	76.0	8.50
2	.hack//Outbreak Part 3	2002.0	0.140000	0.110	0.170000	0.040000	0.460000	70.0	8.20
3	007 Racing	2000.0	0.300000	0.200	0.000000	0.030000	0.530000	51.0	4.90
4	007: Quantum of Solace	2008.0	0.306667	0.225	0.006667	0.113333	0.653333	63.8	5.42
5	007: The World is not Enough	2000.0	0.820000	0.365	0.010000	0.045000	1.235000	61.0	6.70
6	100 Classic Books	2009.0	0.130000	0.520	0.000000	0.020000	0.670000	70.0	6.50
7	101-in-1 Explosive Megamix	2008.0	0.050000	0.130	0.000000	0.020000	0.200000	46.0	0.00
8	101-in-1 Sports Party Megamix	2010.0	0.020000	0.000	0.000000	0.000000	0.030000	41.0	0.00
9	1701 A.D.	2006.0	0.000000	0.250	0.000000	0.050000	0.300000	79.0	7.70

Once our team checked for NaN values using the *.isna().sum().sum()* command seeing that both *df_combined* and *df_merged* no longer contained NaN values, we started our exploration of video game sales and scores.

```
Name            False
Year            False
NA_Sales        False
EU_Sales        False
JP_Sales        False
Other_Sales     False
Global_Sales    False
Critic_Score    False
User_Score      False
dtype: bool
```

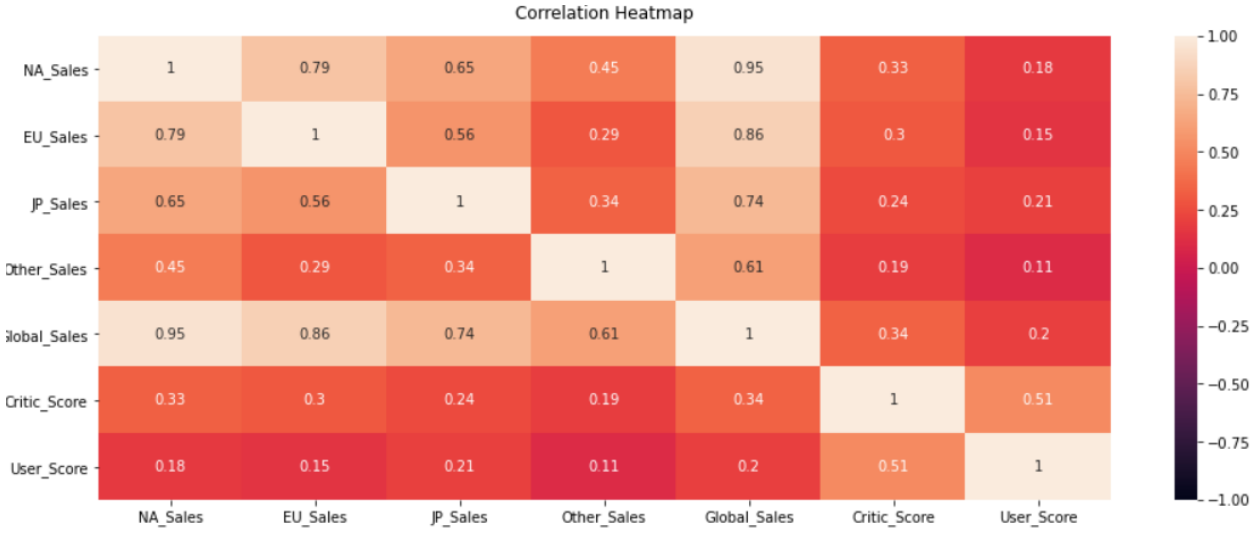
Data Visualizations

To better understand the data we would later visualize, we randomly sampled 500 rows from the *df_merged* data frame. At this stage, our group focused on finding the correlation between sales figures and scores present in our analysis. We did this by defining a new data frame, called *dfv,* that created a sample of 500 rows in a random state. With the *dfv* data frame, we used the command *.describe()* to find the descriptive statistics of our sales and score variables.

	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	User_Score
count	500.000000	500.000000	500.000000	500.000000	500.000000	500.000000	500.000000
mean	0.353937	0.191815	0.086886	0.081635	0.716538	69.749133	6.847390
std	0.717375	0.445927	0.258903	0.366855	1.479659	13.568280	2.062282
min	0.000000	0.000000	0.000000	0.000000	0.010000	26.500000	0.000000
25%	0.050000	0.020000	0.000000	0.010000	0.108750	61.000000	6.487500
50%	0.120000	0.050000	0.000000	0.020000	0.250000	71.000000	7.537500
75%	0.320000	0.160625	0.043750	0.056875	0.676250	80.000000	8.100000
max	6.850000	5.090000	2.130000	7.530000	14.980000	97.000000	9.200000

We were not surprised by the mean of the total worldwide sales being approximately equal to the addition of all the regional sales. However, the average number of total worldwide sales only reached a little above 700 thousand game copies sold. A poor sales average paired with a mean critic score below 70 and a user score below 7 suggests a strong correlation between global sales and scores.

Our group decided to utilize a seaborn correlation heatmap to better see the impact of each correlation in our dataset.



We found that the correlation between total worldwide sales of games and both scores was relatively weak at 0.34 for critics and 0.2 for regular gamers. The highest correlation our group found to 'Global Sales' was with NA sales (as expected); however, it is clear that all regional sales (including NA, EU, JP, and other regions) total to global sales so they are correlated.

Despite these findings, our team decided to further explore the merged dataset by applying visualization techniques to get a better understanding of the relationship between certain variables.

Data Science Exploration

The broad scope of our research question can be broken down into certain attributes as a game can be defined before or after its release.
Alluding to the factors that help our team examine the most sold games include the genre of said games, whether it is a one-off or part of a franchise, platform influence, publisher, which year was it released within, sales of the game outside of the US, etc.

```
RangeIndex: 4575 entries, 0 to 4574
Data columns (total 9 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   Name          4575 non-null   object
 1   Year          4575 non-null   float64
 2   NA_Sales      4575 non-null   float64
 3   EU_Sales      4575 non-null   float64
 4   JP_Sales      4575 non-null   float64
 5   Other_Sales   4575 non-null   float64
 6   Global_Sales  4575 non-null   float64
 7   Critic_Score  4575 non-null   float64
 8   User_Score    4575 non-null   float64
dtypes: float64(8), object(1)
memory usage: 321.8+ KB
```

What are the best selling games in global e-commerce sales?
When observing the best-selling games of all time, our group used the *.nlargest(3, columns = 'global_sales')* to find that Wii staple games published by Nintendo, on average, were the best-selling games of all time.

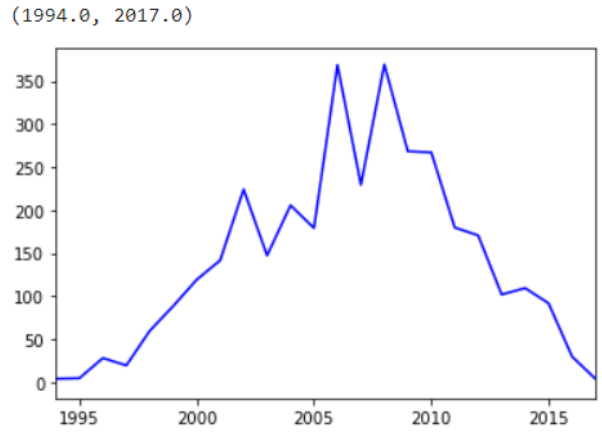
	Name	Year	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	User_Score
4406	Wii Sports	2006.0	41.49	29.02	3.77	8.46	82.74	76.0	8.1
2134	Mario Kart Wii	2008.0	15.85	12.88	3.79	3.31	35.82	82.0	8.4
4408	Wii Sports Resort	2009.0	15.75	11.01	3.28	2.96	33.00	80.0	8.2

Total worldwide e-commerce sales of games by each publisher
To better understand the yearly trend in global sales, our group decided to create a new data frame from the merged dataset called *df_year* grouping by year and select all the regional sales columns to be summed into a total sales figure.

When computing for the years with the largest number of games sold, our group determined that the mid-2000s was when global sales were highest, skyrocketing to around 370 million sales from 2006 to 2008. The majority of that change came from the exponential increase in NA sales in the same period.

Year	NA_Sales	EU_Sales	JP_Sales	Global_Sales
2006.0	179.659893	100.164381	49.214929	368.658869
2008.0	173.862262	105.417643	49.275667	369.289262
2009.0	134.503452	83.432167	21.440333	268.318298

To better understand the yearly trends in games sold, our team decided to graph a line plot using *plt.plot(df_year.Global_Sales, 'b')* where we observed the global sales of every game sold from *df_year*.



We found that sales figures steadily rose from the 1990s into the 2000s, then it rose to over 350 million sales in 2006 and again in 2008 before drastically falling.

Our group believes that the spike in yearly game sales, especially from 2006 to 2010, was heavily influenced by one of the top gaming publishers in our dataset. We decided to create a

new data frame from the combined dataset called *df_pub* grouping by a publisher and computing the sum of global sales by each publisher in our dataset.

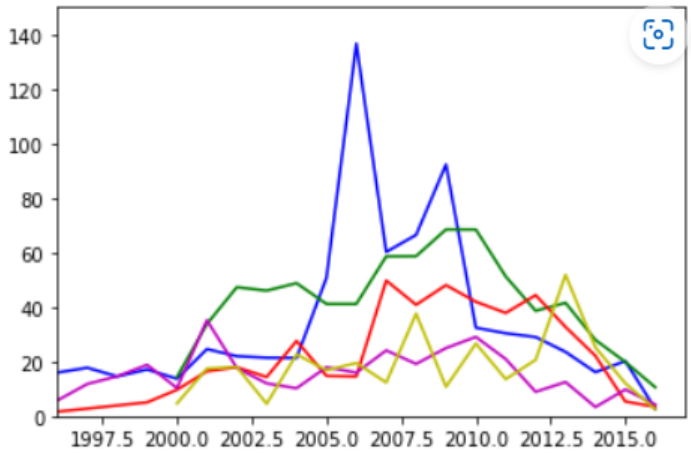
Publisher	Global_Sales
Nintendo	728.72
Electronic Arts	716.16
Activision	447.85
Sony Computer Entertainment	326.62
Take-Two Interactive	316.35

When using the *.nlargest (5, columns = 'global_sales')* function, our group found that the top five publishers are Nintendo at around 728 million dollars, Electronic Arts at around 716 million dollars, then a steep drop for Activision at 447 million, Sony Computer Entertainment at 326 million, and Take-Two Interactive at 316 million.

After we grouped by year and summed the yearly sales for all five gaming publishers, we used the *plt.plot* command to observe their yearly trend of game sales.

```
(1996.0, 2017.0)

(0.0, 150.0)
```

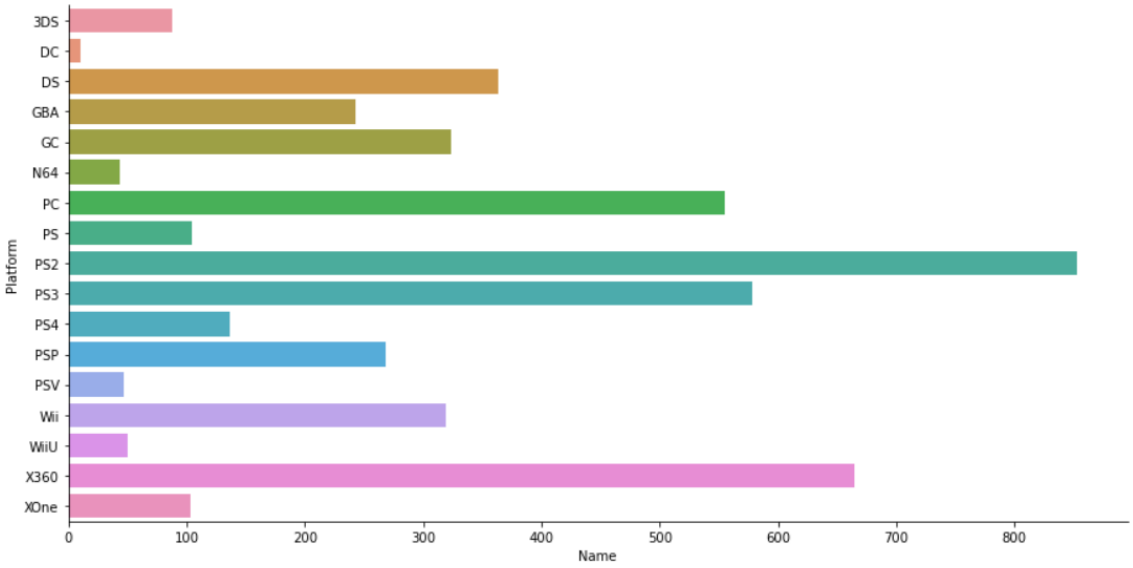


Our group found that Nintendo's total worldwide sales of games after 2005 contributed to the spikes in sales in 2006 and 2008. Outside of Nintendo games, other top publishers had a steady rise in sales.

Total games sold per console on e-commerce platform

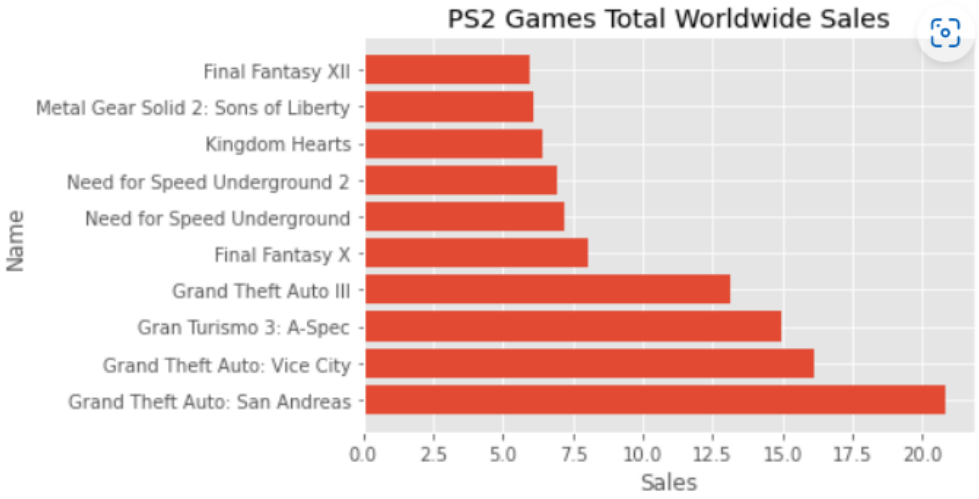
The success of consoles has had a major impact on publisher success so our team wanted to determine which console had sold the most games. A closer examination of the 'Platform' column in *df_combined,* we have found that the top three consoles with the most sold games are the Playstation 2 (853 million), Xbox 360 (665 million), and the Playstation 3 (578 million). Thus, the data implies that gamers ultimately prefer playing games on Playstation consoles over every other console.

We used the *catplot* function to display a bar chart that confirms that the top platforms with the most sold games are non-Nintendo consoles; first the Playstation 2, followed by the Xbox 360, and the Playstation 3, with PC games not far behind in our dataset.

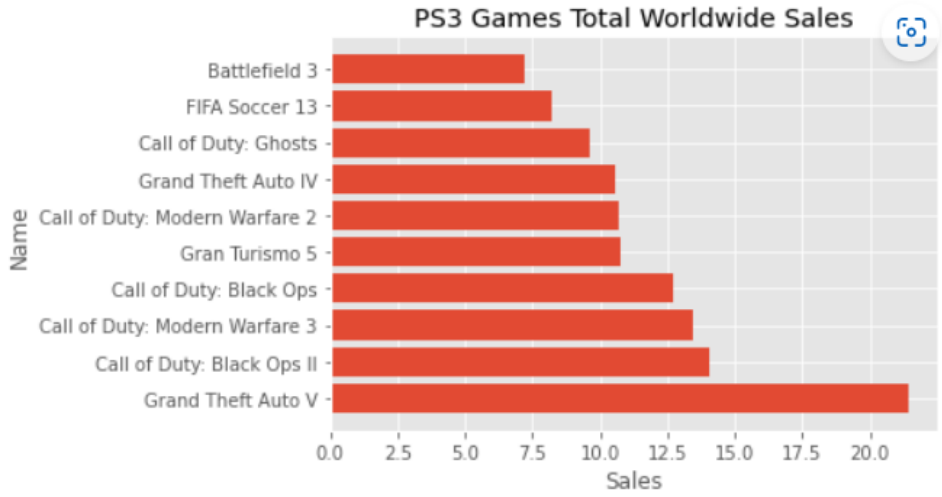


Total worldwide e-commerce sales of games from top consoles
From our previous visualization of total games sold per console, our group found that the best-selling consoles were largely dominated by modern gaming systems including the PlayStation 2 (*PS2*), PlayStation 3 (*PS3*), and Xbox 360 (*X360*) having the most games released.

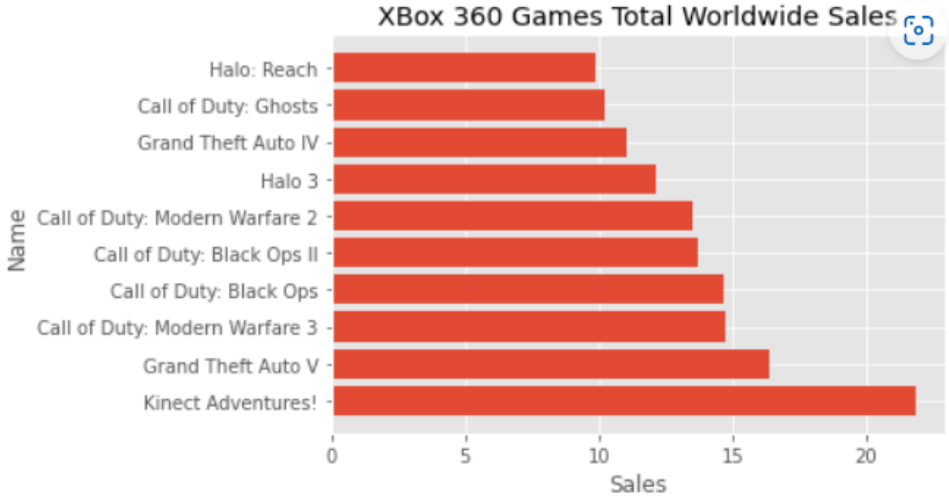
We further wanted to find the top-selling modern games of these platforms using *df_combined* data frame specifying the value *PS2* for the platform column grouping by 'Name' to calculate the aggregate mean of global sales. When using the *.nlargest(10, columns='Global_Sales')* command, the Grand Theft Auto franchise of games before the PS3 remains the four highest-sold games on the PS2 followed by Final Fantasy X.



Our group conducted similar aggregate techniques for the PS3 finding that Grand Theft Auto V remained the most sold game, but saw more representation from the Call of Duty franchise of games.



Finally, when evaluating the Xbox 360 top-selling games, we found it interesting that the best game sold was Kinect Adventures even though the Kinect was largely considered a failure by gamers and critics alike.

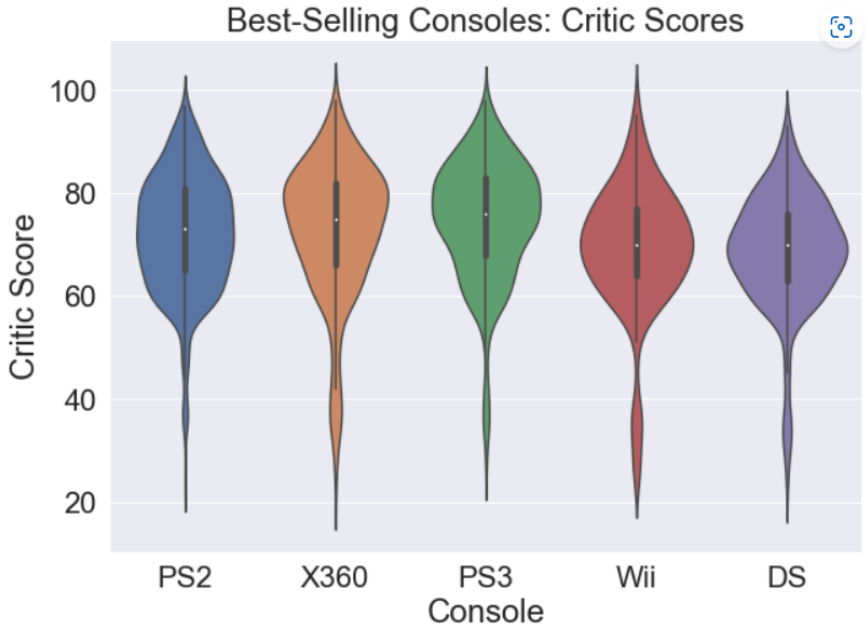


Using visualization techniques helped us observe the true difference in sales figures between these top-selling consoles. When defining name and sales into separate variables than using the *plt.barh(name_'console', sales_'console')* command, we observed that the Xbox 360 had the highest selling games on average, which suggests the overall better performance of the console despite it being ranked third in most games released.

***Critic and user preferences for games released on top consoles**

Critic Score versus Console

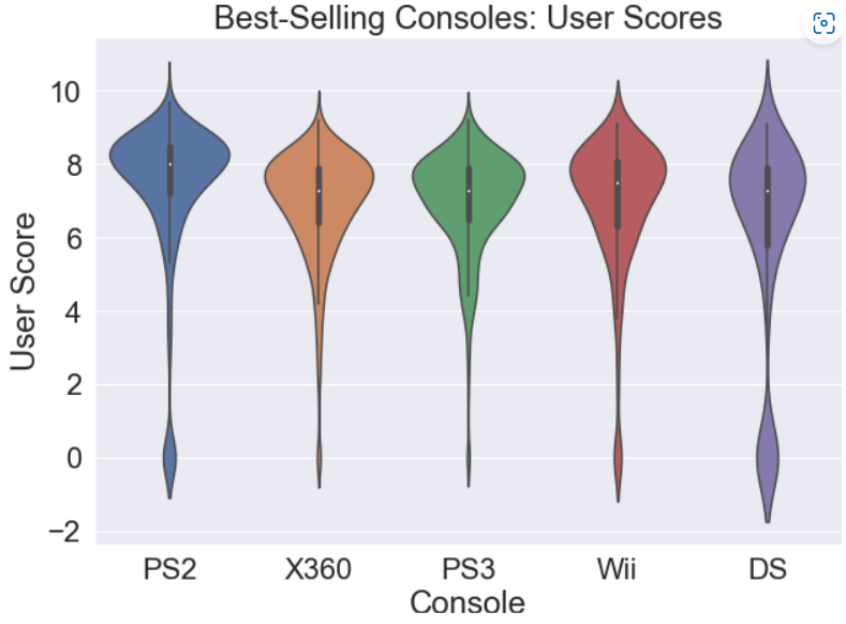
Our group created a violin plot of the best-selling consoles against critic scores including Nintendo consoles Wii and DS.



We found that the PS3 has the highest median score followed closely by the Xbox 360 and PS2. Each console shared a similar data distribution wider near their median score suggesting that the weight of critic scores is highly concentrated around the median value. Nintendo consoles followed a similar distribution but with a higher concentration of scores close to the median.

User Score versus Console

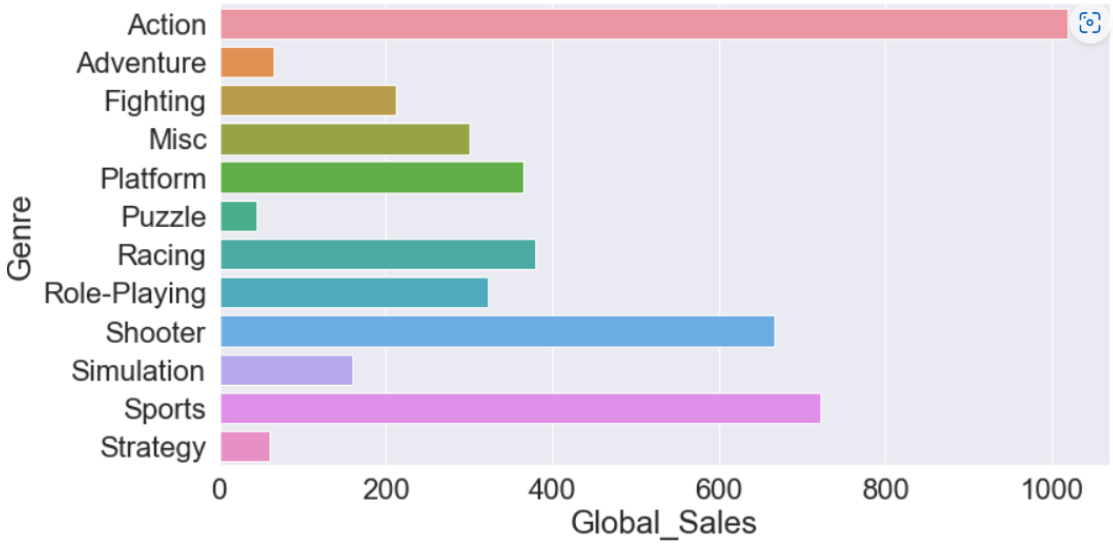
The best-selling consoles against user scores are the PS2 followed by the Xbox 360, PS3, and Nintendo consoles.



In these violin plots, the smaller scale of user scores for the list of games in our dataset led our data distributions to be more concentrated at the median value. The greater spread of critic scores on the top consoles held more weight than user scores, specifically the weight of scores on consoles.

Total worldwide sales figures per genre

In the *df_combined* data frame, we utilized the group by function to filter out the data by genre and the number of games sold globally. The data table revealed that the genre of the game determines what regular gamers prefer based on the highest number of games sold. The data from the table was plotted on a bar graph, using the *catplot* function, and the *.nlargest(3, columns = 'Global_Sales')* command was applied.



We realized that in the top three video game genres that have the highest sales worldwide, the 'Action' genre had the largest total worldwide sales by far with over 1 billion games sold. This number was followed by the 'Sports' and 'Shooter' genres in our dataset.

Critic and user preferences for the genre(s) of games

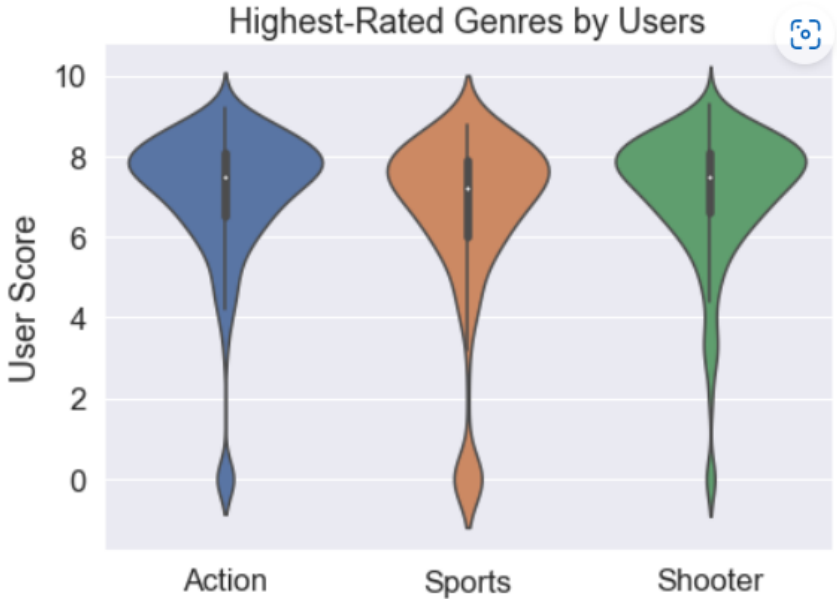
To observe the yearly trend in the genre of games by critics and regular gamers, our group decided to visualize it in a categorical bar plot.

Our team created a new data frame called *df_genre* which used *df_combined* to group by genre and year, then computing the mean using the aggregate function to obtain the mean critic score for each genre. Our group also created four bins that each spanned an average of 4 to 6 years to see a more comprehensive visual.

When evaluating the genre of games against user score, there was relatively no change occurring throughout the four binned years, suggesting consistent criteria and preference of critics when it comes to scoring video games.

User Score versus Top Genres

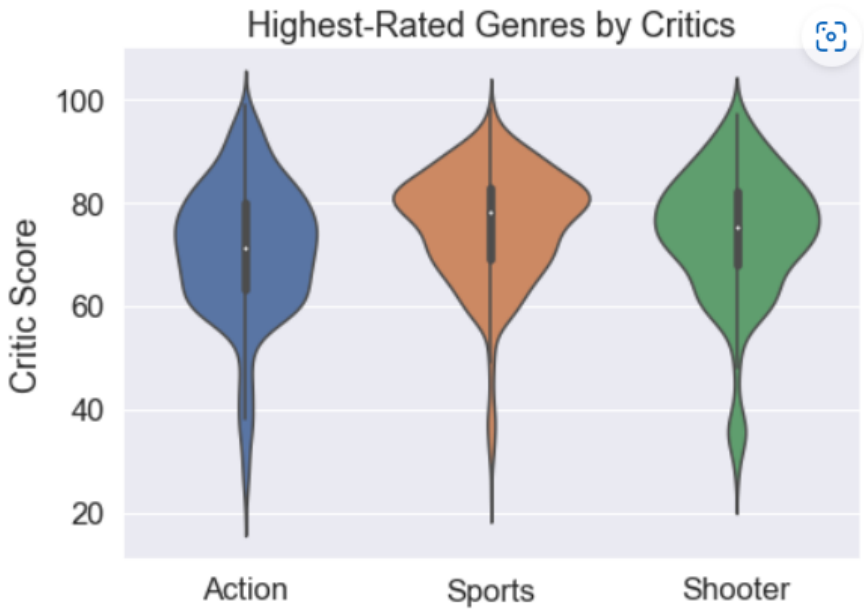
The yearly trends in the genre of games by regular gamers tends to fluctuate much more in our categorical plot.



Out of the top 3 genres, ‘Sports’ and ‘Shooter’ games decreased in average score suggesting that regular gamers have a greater preference for specific genres as gaming interests change over time. Our group believes that as newer games shift to a more popular selling genre, their score for those games shifts. In our plot, we can see the effect in the ‘Adventure’ and ‘Racing’ games as their higher average score steadily rises. Overall, our categorical bar plot reveals that regular gamers hold more opinions about a genre of a game as newer games are released, which could suggest the rise and demise of certain games more than critics may suggest.

Critic Score versus Top Genres

Our group created violin plots of the top genres according to critic scores, specifically, ‘Action,’ ‘Sports,’ and ‘Shooter’ games, to see the spread of the highest-rated genres by regular gamers.



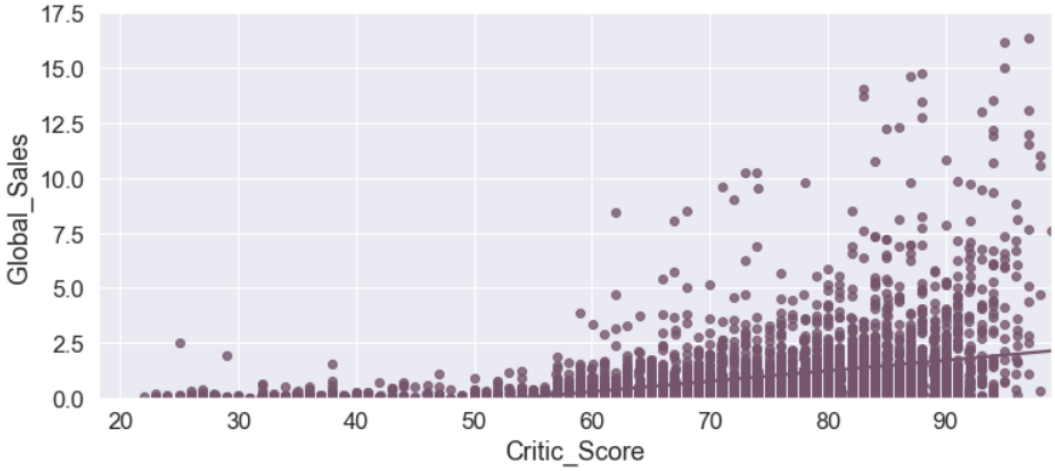
When evaluating our violin plots, ‘Action’ and ‘Shooter’ games have the highest user scores just above scores given to ‘Sports’ games. In these 3 violin plots, we see a similar data distribution that has extremely thin ends and is wider near the median. This suggests that the weight of user scores is highly concentrated around the median value. On the other hand, the highest-rated genres by critics, on average, are ‘Sports’ games. The greater spread of these 3 genres' data distribution means that critic scores given to a genre hold less weight and thus are less likely to draw closer to the median critic score.

Scores on Global E-Commerce Sales

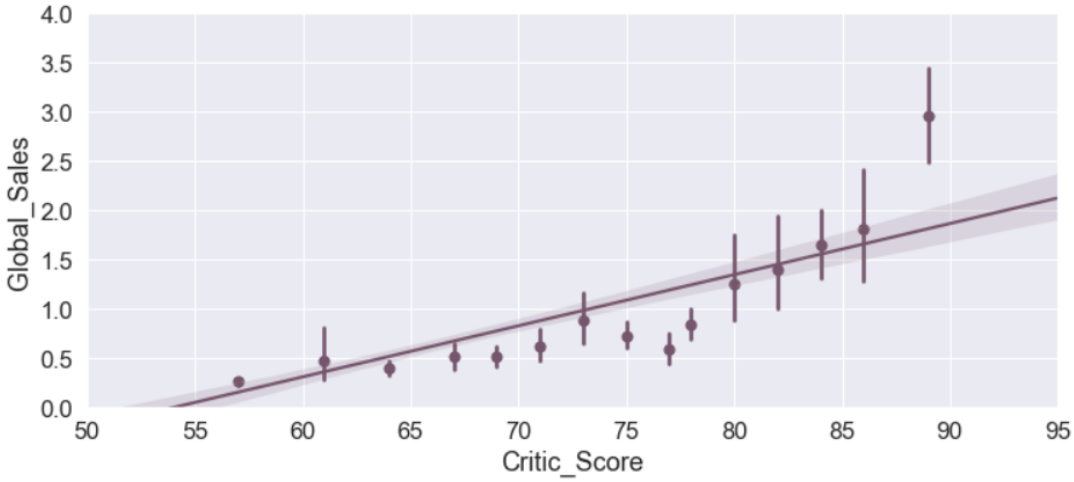
Metacritic score on total worldwide e-commerce sales

Our team proceeded to observe critics' and regular gamers' preferences for consoles and genres. To answer our overarching research question, we brought our attention back to gamers’ scores on sales figures.

Regression plot without bins



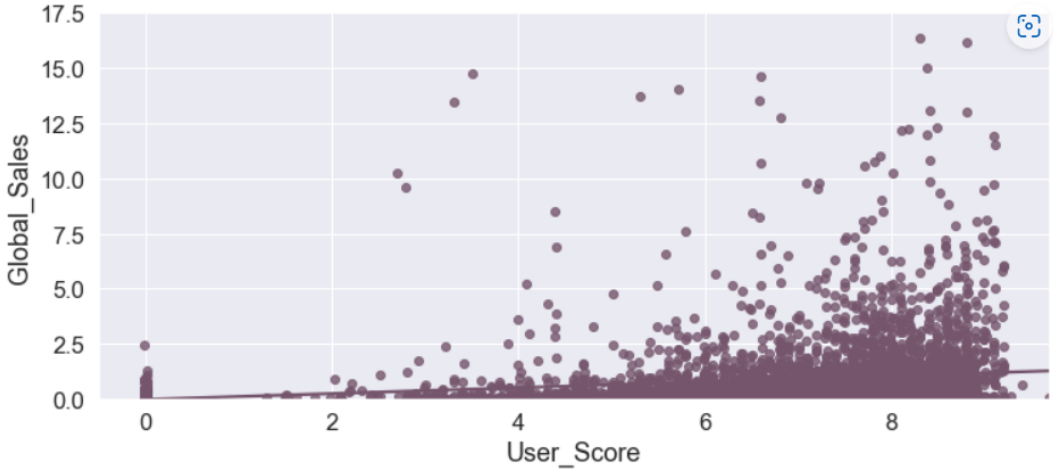
Regression plot with bins



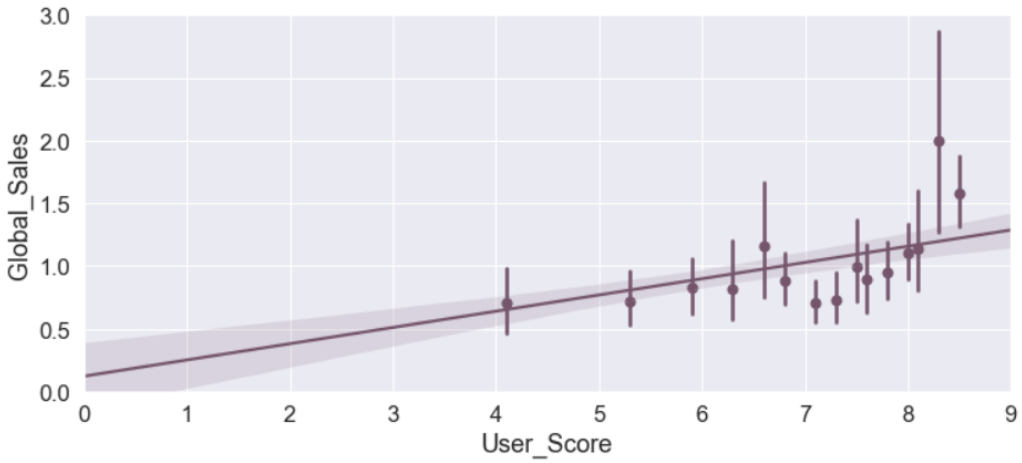
When examining the effect of Metacritic scores on total worldwide sales, we see that critic scores and global sales have a clear positive correlation with each other. Also, as we evaluate critic scores on global sales in a subset of years, we can see that the slope, in the graphic, gets steeper past 80.

User score on total worldwide e-commerce sales
Looking toward the effect of user scores on total worldwide sales, we evaluated user scores on global sales in the whole dataset and we found a less positive correlation between the two variables.

Regression plot without bins



Regression plot with bins



When evaluating user scores in a subset of years, we discovered in the graphic that the slope slightly lowers past 6. Once a game gets a high user score, every additional point has a lower impact on global sales.

Clustering Models

Feature Engineering

When developing a clustering analysis, our group's goal was to identify clusters of games published by Nintendo - which has the most global sales of games - and their NA sales from our merged *df_combined* data frame.

Publisher	Global_Sales
Nintendo	728.72
Electronic Arts	716.16
Activision	447.85

First, we used the *pd.get_dummies(columns=['Publisher'], data=df)* to insert binary values for each row of games in our list. We also cleaned the transformed data frame by dropping the other non-numeric columns: *Name, Platform, Genre,* and *Summary*.

*Apply the KMeans clustering algorithm to identify 2 clusters of games

Generate summary statistics describing the identified clusters' characteristics

When we evaluated our generated summary info, we identified cluster 1 with more sold games in NA at 63.6 percent in NA with 6 percent of those games published by Nintendo than cluster 0 with only 21.9% of sold games with 3.2 percent of those games published by Nintendo.

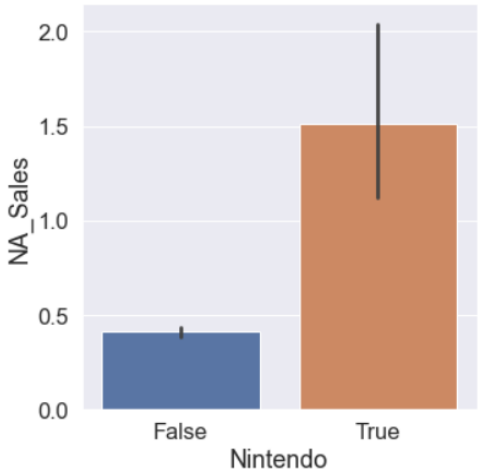
The clustering analysis identified a cluster of Nintendo-released games (cluster 1) that are sold more in North America (NA) than the other cluster in 0.

cluster	Year	NA_Sales	EU_Sales	JP_Sales	Other_Sales	Global_Sales	Critic_Score	User_Score
0	2007.504295	0.635669	0.391550	0.086815	0.136174	1.255251	80.894775	7.596779
1	2007.175779	0.219004	0.125232	0.024716	0.043975	0.414752	62.023505	6.050996

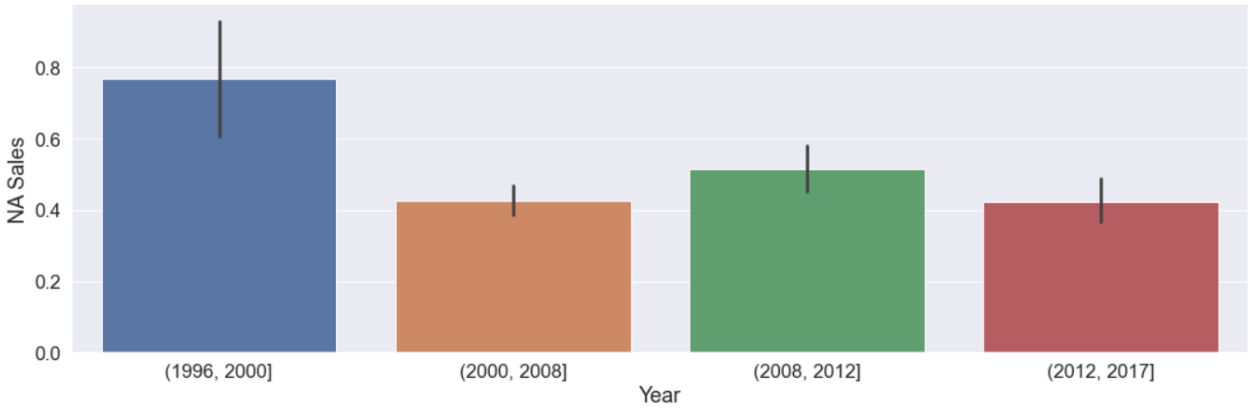
Investigate a finding of the clustering analysis - the propensity of a Nintendo-published game to be sold in NA.

Investigating a finding of the clustering analysis, our group focused on the propensity of a Nintendo-published game to be sold in NA. Nintendo-published games only made up about 10 percent of our list of games so we created a new column called Nintendo setting the condition *[df_combined].Publisher_Nintendo > 0.05* separates the two clusters in our visuals.

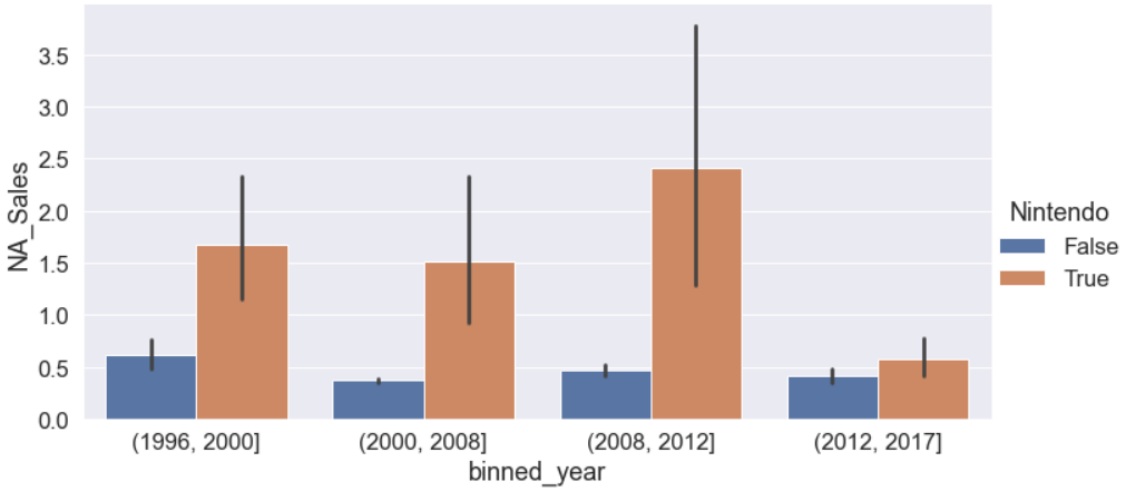
In our clustering analysis, we chose to investigate Nintendo game sales in NA using visualization techniques such as categorical bar plots.



In our first bar plot, we see that there is a strong natural tendency for their games to be sold in NA with an average of 1.5 million Nintendo game copies sold compared to less than 500,000 non-Nintendo game copies sold.



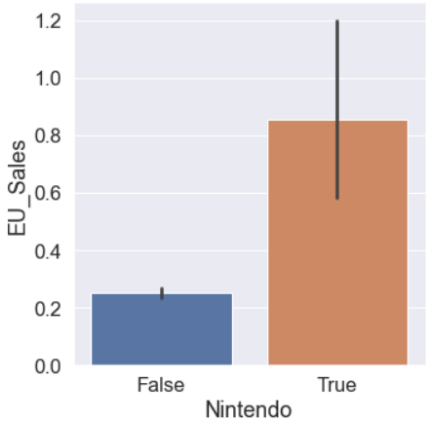
To quantify the effect of Nintendo games over the years, we binned the year's column once again to see that NA sales of the games on our list peaked in the first bin, but fell close to half a million in sales after 2000 and remained consistent since then.



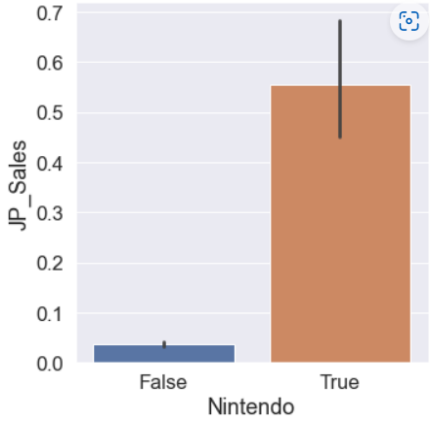
Ordering the plot with the categorical variable Nintendo, we re-evaluated the sales figure of Nintendo games and found that their games are selling, on average, much higher than non-Nintendo games in the NA region until 2012 peaking at over 2 million sales.

Our group's first clustering analysis focused on NA sales of Nintendo games, but our group wanted to compare the NA video game market to other big markets, specifically Europe and Japan.

Propensity of a Nintendo published game to be sold in Europe

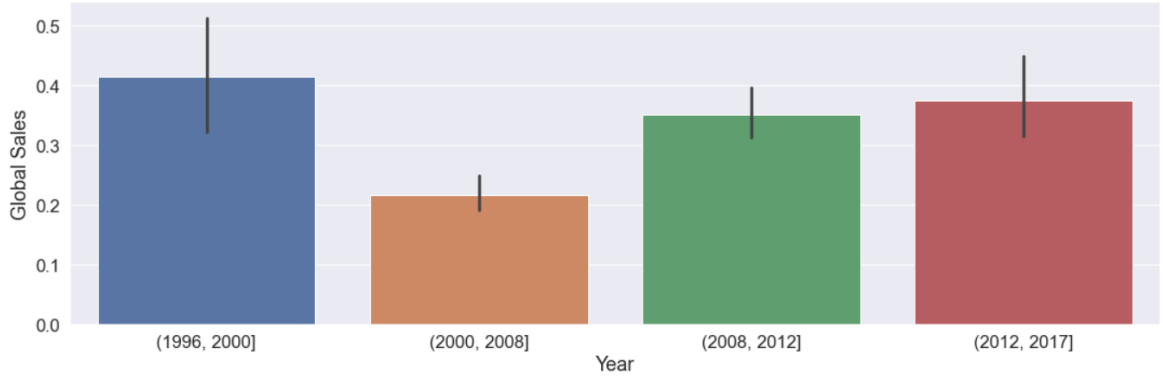


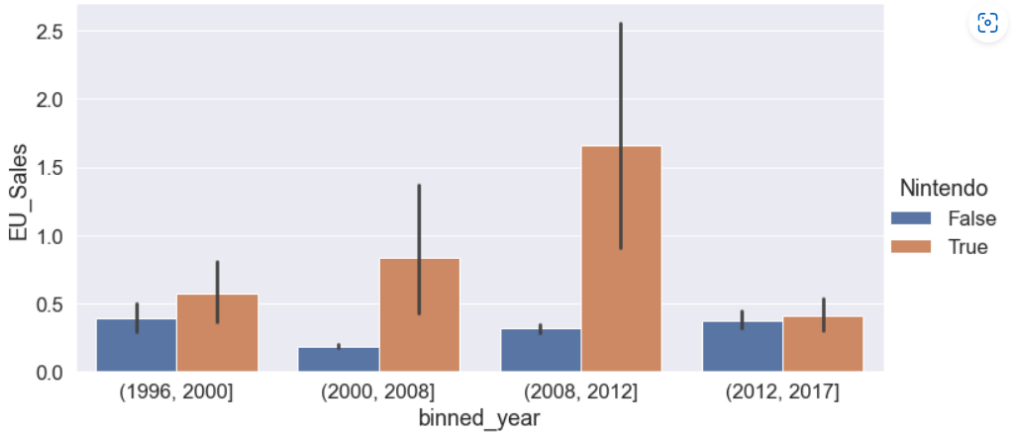
Propensity of a Nintendo published game to be sold in Japan



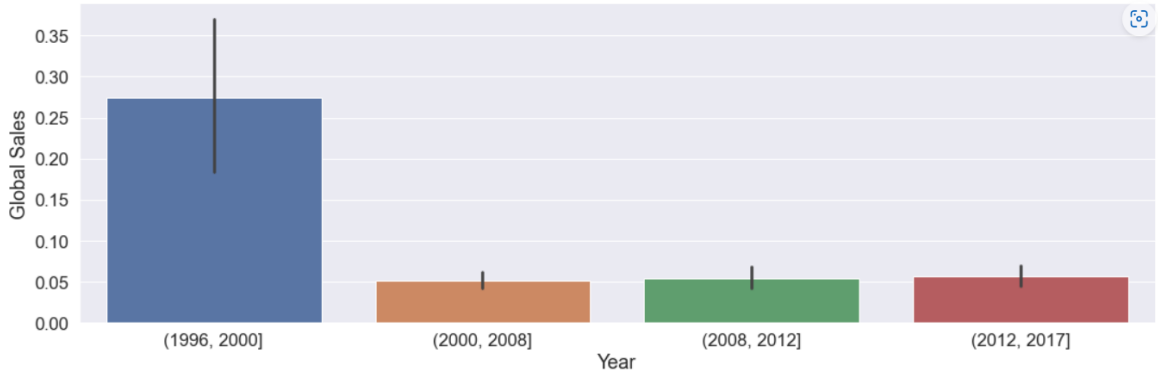
Using similar visualization techniques from our first cluster analysis, we found that European sales of Nintendo game copies, on average, were less than 1 million while Japanese sales of Nintendo games are even lower at less than 600,000 games copies.

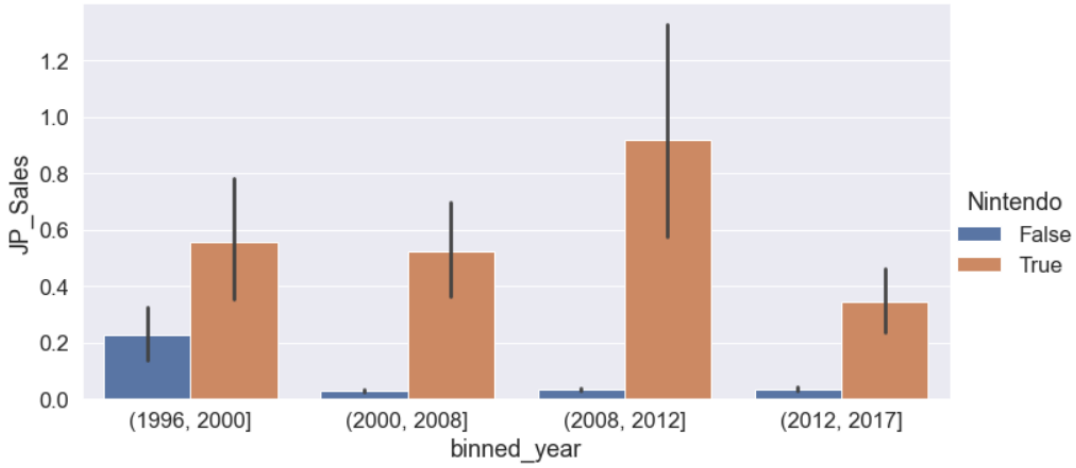
Timeline of Total Number of Games versus Nintendo Games Sold in Europe





Timeline of Total Number of Games versus Nintendo Games Sold in Japan





Both regions' total sales of Nintendo games being less than NA also reflect in our categorical bar plot where Europe's peak sales of Nintendo games only reached a difference, on average, around a million copies compared to non-Nintendo games.

To better quantify our results, our team used the *df_combined* data frame to group by *Nintendo* and computing size and mean of NA, EU, and JP sales.

	NA_Sales	EU_Sales	JP_Sales
	size	mean	mean	mean
Nintendo				
False	4519	0.410188	0.252416	0.035931
True	232	1.512974	0.855172	0.554138

Comparing the sales of all games in our list, the sample size is much smaller for published Nintendo games at 5.13 percent (=232 / 4519). The NA region remains the largest market for the foreign video game publisher Nintendo, even more than the European and domestic Japanese markets.

Comparing clustering methods based upon silhouette score

```
k = 2
KMean with k = 2: 0.5471
Agglo with k = 2: 0.5043
k = 3
KMean with k = 3: 0.4963
Agglo with k = 3: 0.4834
k = 4
KMean with k = 4: 0.4622
Agglo with k = 4: 0.3901
k = 5
KMean with k = 5: 0.433
Agglo with k = 5: 0.3804
*********************************************
Best algorithm is... KMeans  with k = 2
**********************
With Silhouette Score 0.5471147943408368
```

Based on the silhouette score, we also determined that the best algorithm to use in our analysis was indeed 2 clusters with a silhouette score of 0.54. Because of this, our team believes that our mean clusters are well separated from each other and clearly distinguish the regional sales of Nintendo games.

Conclusion

Our team determined that scores and sales of video games are highly correlated and other factors associated with video games link back to their success. In our exploration, we were able to observe the different relationships to global sales of video games and in the end were able to conduct a clustering analysis that evaluated regional sales of video games from Nintendo, one of the largest video game publishers in the world. Our team recognizes that our analysis could be improved by utilizing certain attributes such as score clusters instead of regional sales of Nintendo games. In addition, we missed the opportunity to possibly incorporate a Twitch dataset filled with audience count and demographic data on video game streams, which contributes to the popularity of said video games. Despite this, we were still successful in examining the most sold games from a global perspective.